

A SkinCare projekt

Somfai Ellák - ELTE IK Mesterséges Intelligencia Tanszék

Lőrincz András, Pólik Marcell, Baffy Benjámin,
Szijártó Ádám, Antika Das, Kristian Fenech, Ulbert Attila

Ipari együttműködés: Degetel (Franciaország)
EIT Digital projekt - SkinCare

Orvosi partner: Semmelweis Egyetem, Bőr-, Nemikórtani és Bőronkológiai Klinika
Hosszú Rita, Korózs Dorina, Sárdy Miklós



PROGRAM
FINANCED FROM
THE NRDI FUND

Célkitűzés:

melanoma felismerés dermoszkópos ill. makro fényképfelvételek alapján

Motiváció:

- *melanoma*: bőr pigmentsejtjeinek rosszindulatú burjánzása
- legagresszívebb bőrrák típus: bőr daganatos megbetegedéseinek csak 5-10%-a, mégis a legtöbb bőrrákból eredő halálesetet okozza
- diagnosztizált esetek száma az utóbbi évtizedekben folyamatosan növekszik „évtizedenként duplázódik”
- korai diagnózis: 99% 5-éves túlélés
késői diagnózis: 50%-ban végzetes áttéteket okoz

*korai diagnózis elősegítése
gépi tanulási módszerek segítségével*

Melanoma vagy anyajegy (nevus) ?



Melanoma



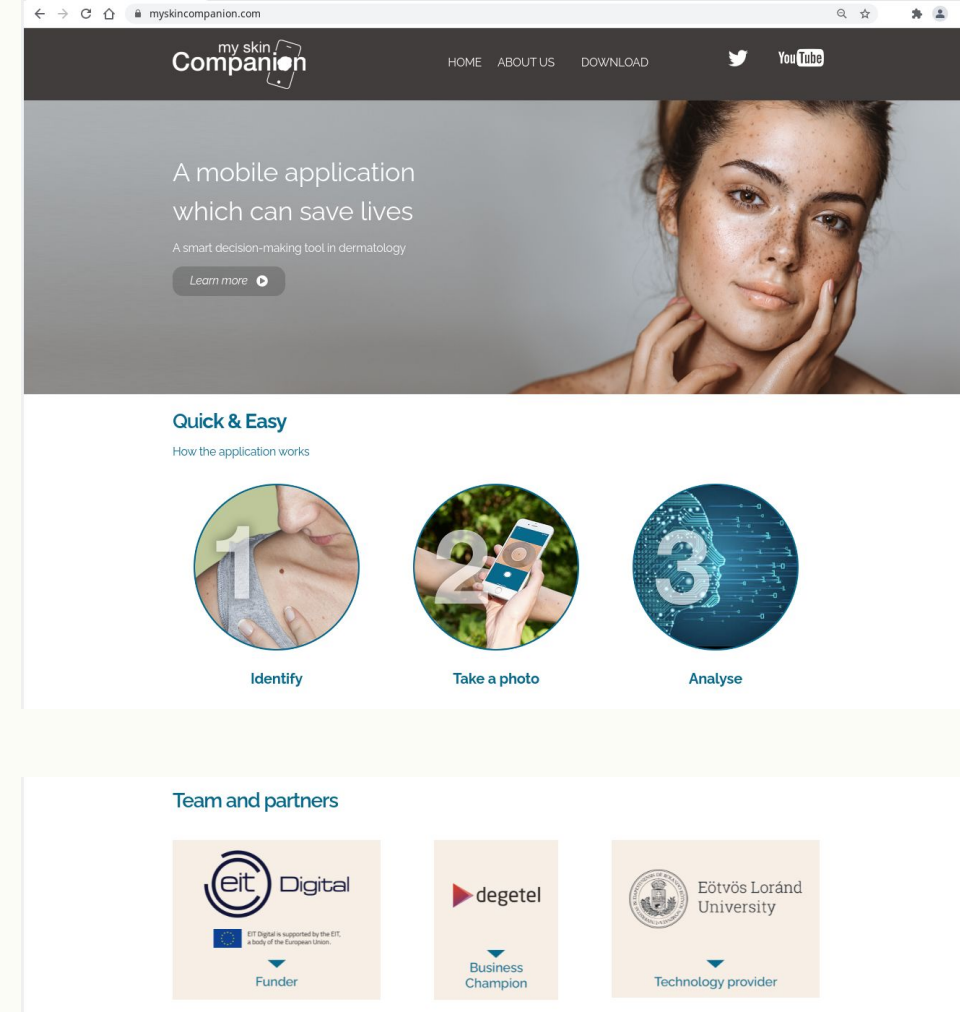
anyajegy

Ipari projekt

EIT Digital / SkinCare projekt

Mobiltelefon alapú alkalmazás
automatizált melanoma detektálásra.

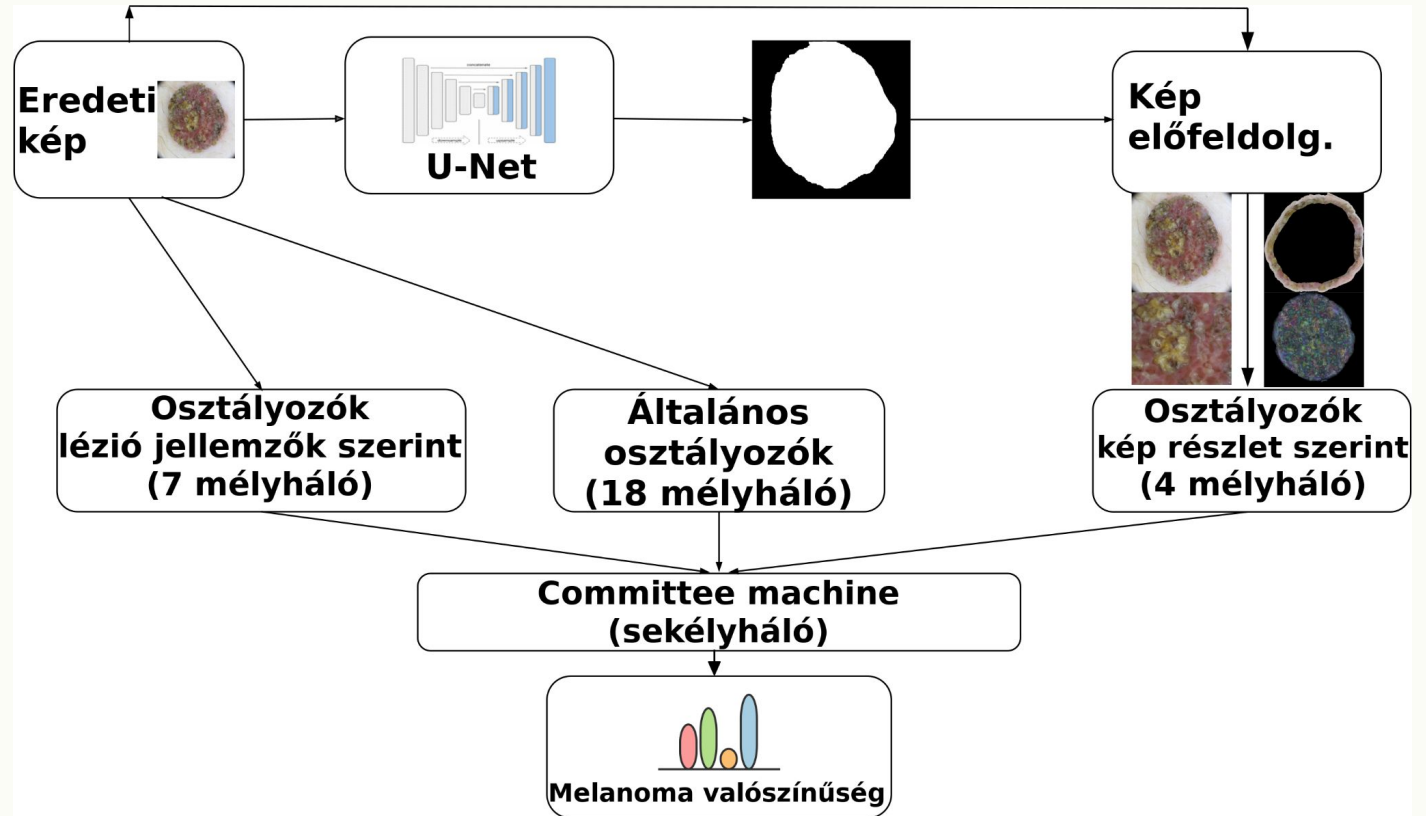
ELTE feladata: döntési algoritmus kidolgozása
és implementálása (felhőben fut).



Módszerek:

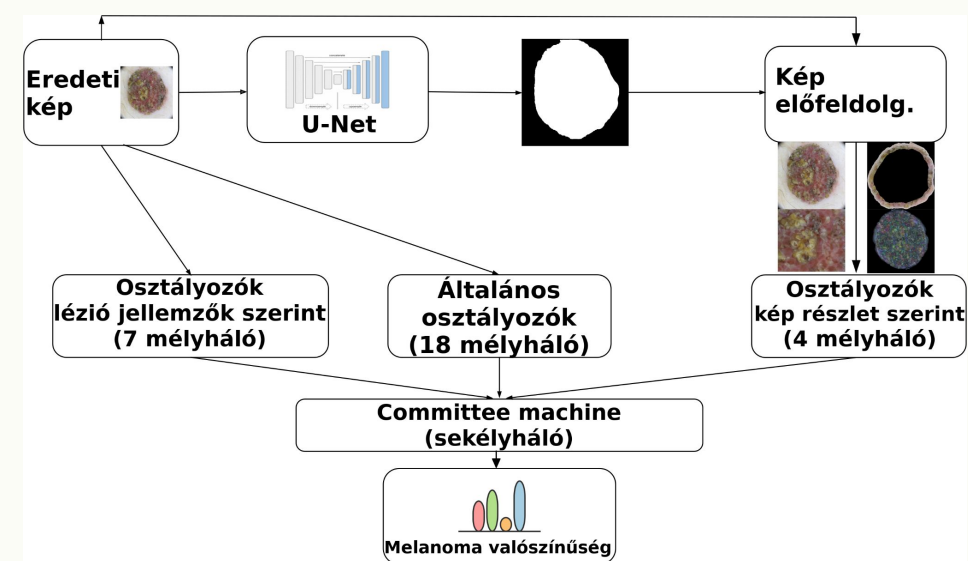
bőrelváltozás (lézió) osztályozása:

7+18+4 mély neuronháló modell kombinációja



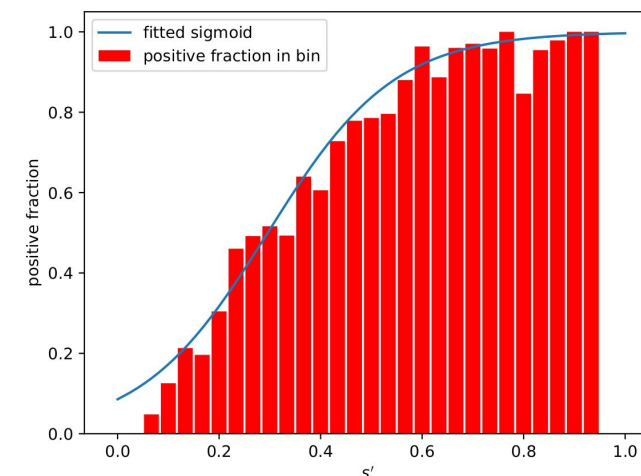
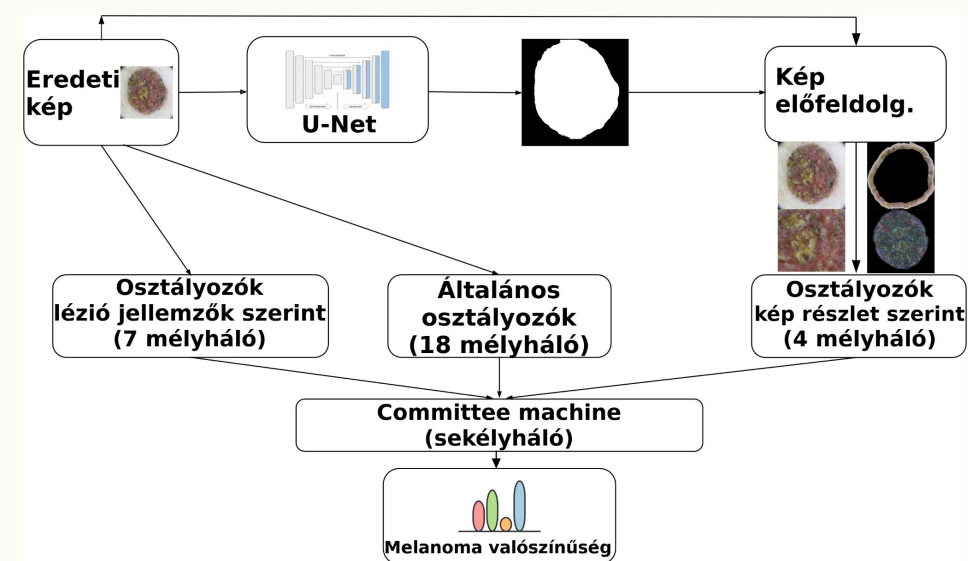
Módszerek:

- „nyers erő” osztályozás (18 db)
erős - nagyon erős mélyhálók (EfficientNet b7-ig)
2020 SIIM-ISIC-Kaggle challenge győztese
- osztályozási cél: karakterisztikus lézió jellemzők [features] (7 db)
pl. kék-fehér fátyol, pigment hálózat, pontok és szemcsék, stb.
közepes erősségű mélyháló (EfficientNet b4)
- képi előfeldolgozáson alapuló jellemzők (3 db)
lézió szegmentálás, kép részlet kiemelés:
lézió közepe (apró részletek), lézió széle, szín-aszimmetria
közepes erősségű mélyháló (EfficientNet b4)



Módszerek:

- eredmények kombinálása:
tanult módon - „committee machine”
sekély háló (feed forward, 3 rejtett réteg)
- kalibrált valószínűség:
kimenet gyakran nem értelmezhető közvetlenül
kalibrálás: nyers kimenő softmax értéket
(előfeldolgozás után) dobozokba gyűjtjük,
dobozon belül pozitív ráta
(„reliability diagram” a valószínűségi meteorológiai
előrejelzésben)



Projekt - EIT

Feladat: algoritmus fejlesztése és implementálása mobiltelefonnal készített bőrlézió-képek melanoma / nem melanoma osztályozására.

Kihívás: normál fényképek („makró” vagy „klinikai”), nem dermoszkópos felvételek

- kevésbé mutatja a karakterisztikus jellemzőket (bőrfelület tükröződik), mint a dermoszkópos felvételek
- sokkal kevesebb nyilvánosan hozzáférhető képi adatbázis

Megoldás:

- transfer learning - előtanulás dermoszkópos képekre, áttanulás makró képekre
technikai megvalósítás: több lehetőség lehetséges. Legjobb: mélyhálók csak dermoszkópos képekre tanítva, committee machine: makró képek eredményeire tanítva

Derm7pt (tanító halmaz): 73%

Med-node: 84%

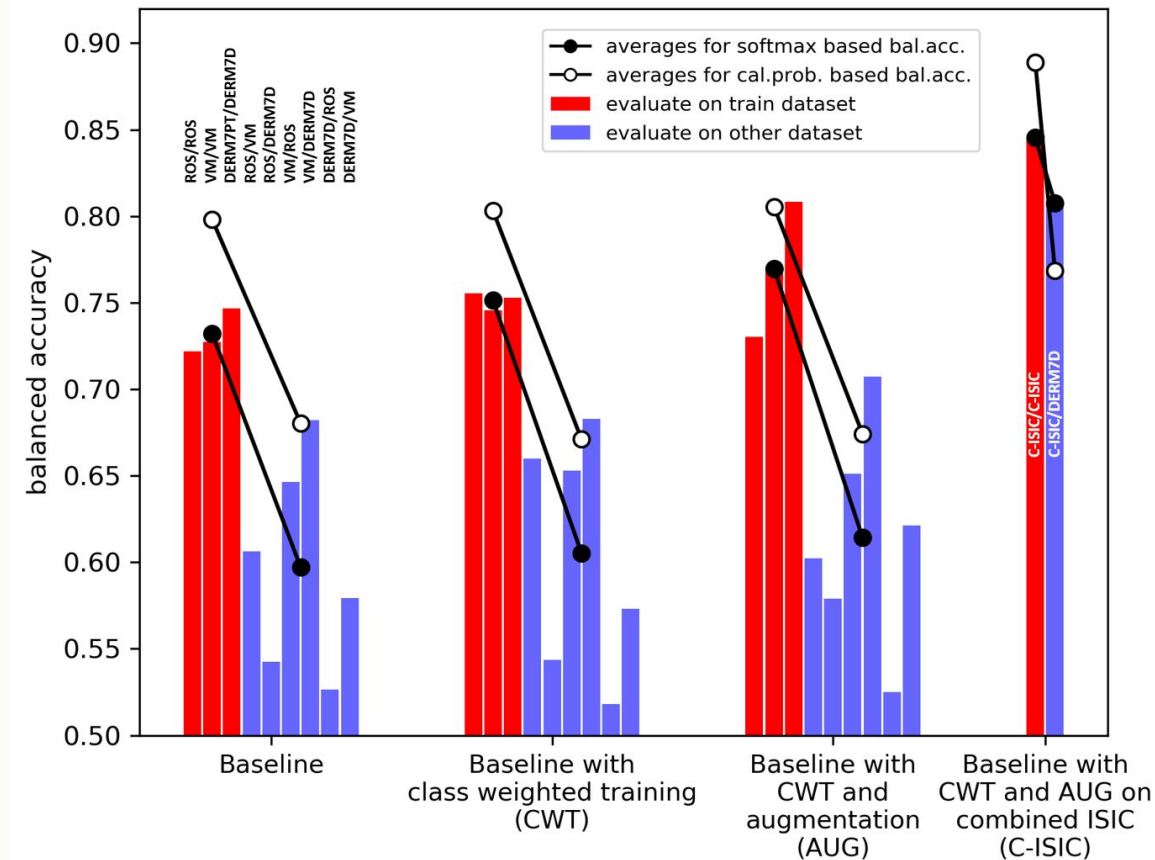
Semmelweis-makró: 73%

Projekt - Adathalmaz-függőség (csak dermoszkópos képek)

- egyik első tapasztalat: nyilvános dermoszkópos adathalmazon 70-80%-osra (multiclass balanced accuracy) betanított modell a Semmelweis-dermo adathalmazon csak 40%-ra teljesített
- általánosan: az A adathalmazon betanított modell az A adathalmaz teszthalmazán (vagy kereszt-validációban) jobban teljesít mint a B adathalmazon, és fordítva is: a B -n betanított modell jobban teljesít a B -n mint az A -n.
Pedig: A és B ugyanabból az eloszlásból kellene mintát vegyen (csak: másik klinika)
- feladat: milyen módon csökkenthető az adathalmaz-függőség?
teljesítmény javítása új, ismeretlen forrásból származó adatokra

Projekt - Adathalmaz-függőség (csak dermoszkópos képek)

- baseline:
1 db közepes mélyháló
- változatok:
 - osztály-súlyozás
 - erős augmentálás
 - nagy adathalmazra tanítva
- kalibrált valószínűség:
kezdetben segít, de okozhat túltanulást



Projekt - Adathalmaz-függőség (csak dermoszkópos képek)

Balanced accuracy		predict on training dataset		predict on other dataset		
		C-ISIC	DERM7D	DERM7D	PH2	Semmelweis
single models	C-ISIC Baseline	0.8454		0.808	0.822	0.815
	C-ISIC Calib.Prob	0.8888		0.769	0.778	0.823
ensemble models	AVG-18	0.8912		0.792	0.841	0.770
	Calib.Prob.AVG-18	0.9345		0.825	0.919	0.818
	Committee-18	0.9344		0.836	0.909	0.832
	Committee-22	0.9356		0.848	0.909	0.831
	Committee-29	0.9351	0.848		0.912	0.836
estimated error		±0.0004	±0.002	±0.002	±0.003	±0.004

- sokaság-alapú modellek erősebbek mint az egyedi modellek
- a state-of-the-art (18 nagy teljesítményű modell) javítható néhány, jól kiválasztott, jóval kisebb számításigényű háló hozzáadásával és a mélyháló-kimenetek tanult módon (committee machine) történő kombinálásával

Folyamatban lévő projektek

- **Magyarázhatóság:**

a melanóma-jellemzőket is felhasználó módszerek: többlet információ hasonló bőrgyógyászok által is használt protokollhoz. Lehetséges felhasználás:

- automatizált előszűrés, csak a valószínűsíthetően pozitív és a kétes esetek kerülnek orvoshoz, magyarázattal.
- bőrgyógyász hallgatók oktatása

- **Adathalmaz-függőség makró képekre:**

dermoszkópos -> makró transfer learning szisztematikus vizsgálata

- **Melanóma előrehaladottság becslés:**

erős korreláció a vastagsággal. Osztályozás vastagság szerint 3 osztály

- **Fals negatív ráta kontroll:**

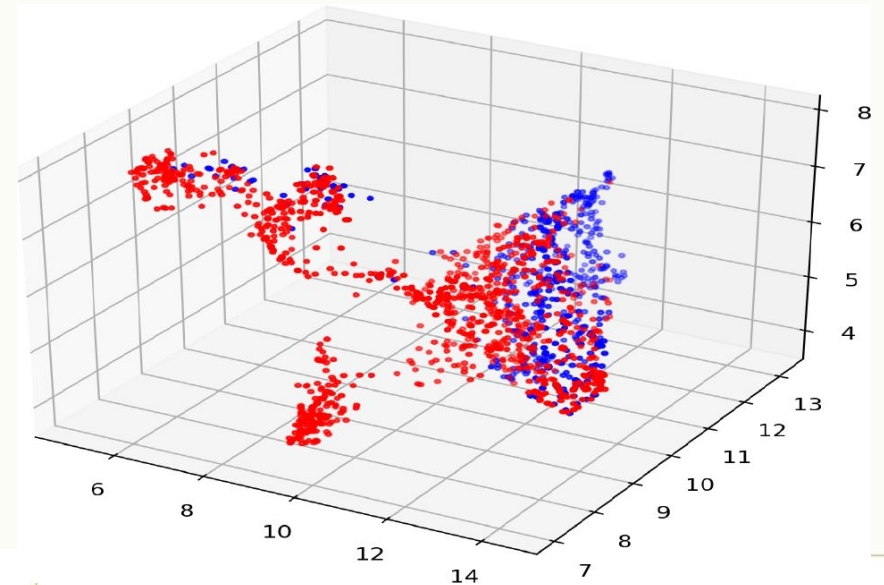
feature map (döntés előtti utolsó CNN réteg)

3D beágyazása UMAP-pel

(nemlineáris dimenzió redukció)

negatívra osztályozott minták: TN v FN

- jobban szétválaszthatóak



Összefoglalás

- 29 mély neuronhálóból álló csoport összeállítása:
meghaladja a 2020 SIIM-ISIC-Kaggle challenge győztes (state-of-the-art) teljesítményét,
csekély többlet számítási kapacitás felhasználásával
(4 adathalmazon tesztelve)
- alapvető lézió jellemzők (features) használata:
magyarázhatóság

Köszönöm a figyelmet!

www.elte.hu



NATIONAL RESEARCH, DEVELOPMENT
AND INNOVATION OFFICE
HUNGARY

PROGRAM
FINANCED FROM
THE NRDI FUND